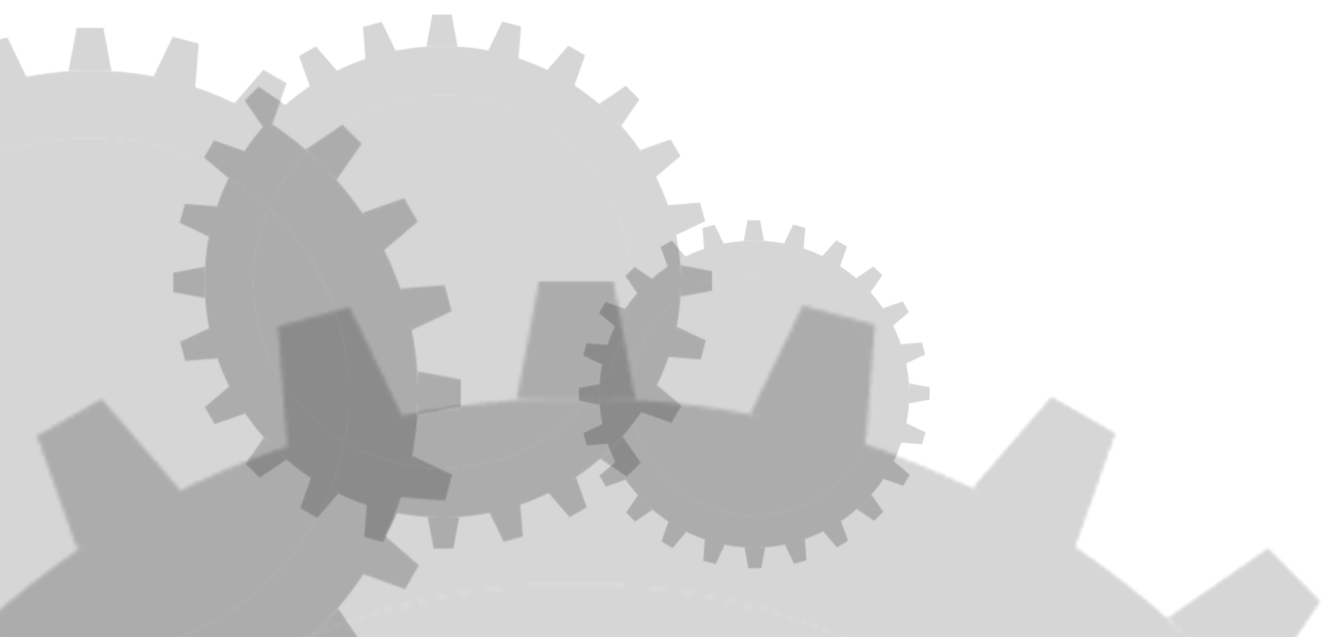
Reference and cognition

Vogelzang, Margreet

4

Reasoning about alternative forms is costly:
The processing of null and overt pronouns in Italian using pupillary responses

This chapter was submitted as:
Vogelzang, M., Foppolo, F., Guasti, M.T., Van Rijn, H., & Hendriks, P. (submitted). *Reasoning about alternative forms is costly: The processing of null and overt pronouns in Italian using pupillary responses*. Manuscript, February 2017.



Abstract

Different words are generally expected to have different meanings. However, many words have seemingly similar meanings. An example are null and overt pronouns in Italian, which both refer to an individual in the discourse. An open question is whether the interpretation of a form is affected by the existence of another form with a similar meaning. Here we show with a pupillary response study that null and overt pronouns are processed differently. Specifically, we found that null pronouns are less costly to process than overt pronouns, and replicated earlier studies indicating that null pronouns tend to refer to the discourse topic and overt pronouns to a different referent than the topic. We explain this difference by arguing that a costly additional reasoning step, reflected in the pupillary response, is required for the processing of marked overt pronouns, but not for unmarked null pronouns. We compared these results to data from a corresponding study on Dutch, a language with overt but no null pronouns, and demonstrate that Italian pronouns are processed differently from Dutch pronouns. These findings suggest that the processing of a marked form in a given language is influenced by alternative forms within that language, making its processing costly.

4.1 Introduction: Reasoning about alternative forms

A fundamental property of language is reference. Words and phrases in a language tend to refer to things, people, thoughts, situations, and events in the world. Processing such referring expressions and understanding what they refer to is essential for successful communication. Here we will discuss how the understanding of referring expressions is infuenced by alternative forms available in the language.

Commonly used referring expressions are personal pronouns such as the English subject pronoun *he*. Not all referring expressions are processed in the same way. For example, processing referential full noun phrases (NPs[2]) such as *the boy* has been found to be less effortful than processing anaphoric pronouns such as *he* (Vogelzang, Hendriks, & Van Rijn, 2016; see also Chapter 3 of this dissertation). Even when these two expressions can refer to the same person, they do not carry the same information. The noun phrase *the boy* carries more explicit information about the intended referent in the discourse than a pronominal form like *he*. The latter, in fact, refers more generally to a male individual present in the discourse context and thus, for example, might also refer to a man in a context in which there are both a boy and a man. Conversely, *the boy* can only refer to a boy.

In general, if there are two forms available in a language that share the same referential space and can be used to express similar meanings, one of them is a briefer, more lexicalized and possibly more frequent form and is *unmarked* (Greenberg, 1966; Horn, 1984). The other form is subsequently a *marked* form. These two forms differ in meaning, with the unmarked form having an unmarked, stereotypical meaning, and the marked form having a less stereotypical meaning for which the unmarked form is usually not used. This is referred to as Horn's division of pragmatic labor (Horn, 1984, 2004). Such a difference is similarly illustrated by Kiparsky's (1982) notion of partial blocking, according to which some forms have a specific meaning, that blocks other, more general, forms from picking up this meaning. For example, the verb *killed* is an unmarked form with a stereotypical, specific, meaning (McCawley, 1978). When the less typical phrase *caused to die* is used, this generally means that X caused Y to die by another means than through killing Y directly, as in the sentence '*X caused Y to die on Sunday by poisoning him on Saturday*' (Fodor, 1970). So, the marked or general form *caused to die* carries a meaning which is different from the stereotypical meaning of the unmarked form *killed*.

Both Horn (1984) and Kiparsky (1982) argue that the meaning of the marked form is derived from the meaning of the unmarked form. This is in line with Grice (1975), who proposed that listeners generally assume speakers to be cooperative, and therefore perform a pragmatic reasoning step called an implicature if the speaker used a form that is not as informative, brief, or relevant as possible. Thus, if a speaker uses a marked form instead of an unmarked form, the listener may reason that the speaker wanted to express a meaning that is different from the meaning expressed by an unmarked form.

[2] Also called determiner phrases (DPs) in the linguistic literature

The specific form used therefore provides additional information about the intended meaning. Blutner (2000) formalized this type of pragmatic reasoning in the linguistic framework of Optimality Theory (Prince & Smolensky, 2004) and argued that, when interpreting a form, alternative forms and their meanings must also be taken into account by optimizing bidirectionally. Although Blutner introduced bidirectional optimization as an algorithm applying to forms and meanings simultaneously, bidirectional optimization can also be conceived of as a serial process. In interpretation, this serial process starts with an optimization step from the perspective of the listener (proceeding from an input form and determining the optimal meaning for that input form), followed by an optimization step from the perspective of the speaker (proceeding from an input meaning and determining the optimal form for that input meaning) (Hendriks, Van Rijn, & Valkenier, 2007; Van Rij, Van Rijn, & Hendriks, 2010). This second, additional, optimization step from meaning to form ensures that the meaning assigned to a given form in the first optimization step is indeed the meaning that the speaker intended to convey by using the given form. Crucially, this second bidirectional optimization step – or pragmatic reasoning step – is only relevant when there are potentially better (i.e., more informative, briefer, more relevant) alternative forms for the given form. If no such alternative forms are available, whether or not the second optimization step is executed is not relevant for interpretation.

In this chapter, we will examine whether the availability of alternative referential forms in a given language affects the meaning and processing of a chosen referential form, because an additional reasoning step has to be performed. This will be tested in Italian, a null subject language which not only allows for full NPs and overt pronouns in subject position, but also for null pronouns. Crucially, overt and null pronouns are known to refer to different entities (cf. e.g., Carminati, 2002). Moreover, null pronouns have a stronger antecedent preference than overt pronouns (Carminati, 2002), which suggests that these two pronouns are processed differently. We will use pupillary responses to investigate whether the processing of overt and null pronouns in Italian is similar or not. Our main hypothesis is that overt and null pronouns are processed differently because reasoning about alternative forms is required for the processing of overt pronouns, but not for null pronouns.

Additionally, a cross–linguistic comparison will be carried out to investigate whether similar forms have different meanings in languages which differ in their availability of alternative forms. To this end, the results from the experimental study in Italian will be compared to the results of the study of Vogelzang et al. (2016; see Chapter 3 of this dissertation) on the processing and interpretation of anaphoric subjects in the non–null subject language Dutch.

4.1.1 Anaphoric subjects in Italian

In non–null subject languages like Dutch and English, a speaker can use a full NP or an overt pronoun (such as *he*) as the subject of a sentence (cf. 4.1). In contrast, in null subject languages like Italian and Spanish, speakers have a third option and can choose

between using a full NP, an overt pronoun (such as *lui* 'he'), or a null pronoun (omitting the subject, '∅', cf. 4.2).

(4.1) a. De jongen eet een appel.
The boy is eating an apple.
b. Hij eet een appel.
He is eating an apple.

(4.2) a. Il ragazzo mangia una mela.
The boy is eating an apple.
b. Lui mangia una mela.
He is eating an apple.
c. ∅ mangia una mela.
∅ is eating an apple.

The meaning of an anaphoric subject depends on the discourse it appears in. Generally, a pronoun (e.g., *he*) refers to an entity that has already been introduced and thus is known in the discourse, that is highly prominent in that discourse context, and that appears in a syntactically prominent position in the sentence (e.g., Ariel, 1990; for an overview see Arnold, 1998). Conversely, a linguistically more constrained form such as a full NP (e.g., *the boy*) refers to a referent that is less prominent (e.g., new to the context, Ariel, 1990; Chafe, 1994; see Figure 4.1). Different terms have been used to describe a referent's prominence in discourse, such as accessibility (Ariel, 1990), givenness (Gundel et al., 1993), and topicality (Givón, 1983). We will use the notion of prominence as an operationalization of a multitude of factors that influence how accessible, given, or topical a referent is.

As can be seen in Figure 4.1 on the left side, overt subject pronouns in non–null subject languages like Dutch generally refer to the most prominent referent (a.o., Ariel, 1990; Givón, 1983; Gundel, Hedberg, & Zacharski, 1993; Vogelzang et al., 2016; see also Chapter 3 of this dissertation), which is often called the discourse topic. The discourse topic is usually realized as the grammatical subject of the previous sentence (Grosz et al., 1995). Subject pronouns specifically often refer to the subject of the previous sentence (Grober, Beardsley, & Caramazza, 1978; Smyth, 1994). Full NPs can be used in non–null subject languages to refer to less prominent referents (a.o., Ariel, 1990; Givón, 1983).

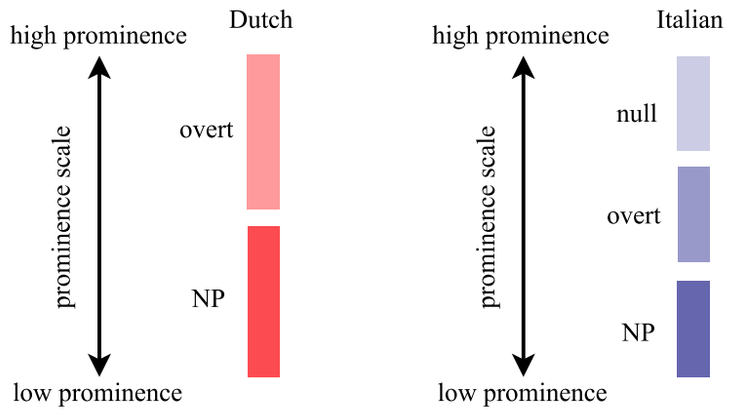


Figure 4.1. A prominence scale indicating that anaphoric subjects that are linguistically more constrained generally refer to referents with a low prominence, and that anaphoric subjects that are less constrained, like pronouns, generally refer to referents with a high prominence (NP: full NP, overt: overt pronoun, null: null pronoun). In the null subject language Italian, an additional anaphoric subject is available compared to Dutch (namely a null pronoun)[3].

As explained, in the null subject language Italian an additional anaphoric subject is available compared to English and Dutch. The availability of this additional null pronoun may cause a different mapping of forms on the prominence scale (Figure 4.1, right side). In agreement with the assumption that different forms have different meanings, Carminati (2002) found clear antecedent preferences for pronouns in preverbal subject positions in Italian: Null pronouns prefer an antecedent in subject position and are therefore read faster and understood better when referring to a subject antecedent. This was shown in sentences like (4.3) (from Carminati, 2002, p.33):

(4.3) a. Quando Mario ha telefonato a Giovanni$_i$, lui$_i$ aveva appena finito di mangiare.
When Mario called Giovanni$_i$, he$_i$ had just finished eating.
b. Quando Mario$_i$ ha telefonato a Giovanni, ∅$_i$ aveva appena finito di mangiare.
When Mario$_i$ called Giovanni, ∅$_i$ had just finished eating.

In Example (4.3), co–reference between two expressions is indicated by the index i. In (4.3a) the overt subject pronoun *lui* must be interpreted as referring to the non–subject *Giovanni.* In (4.3b) the null pronoun (represented by ∅) must be interpreted as referring to the previous subject *Mario.* Therefore, these sentences illustrate that different anaphoric subjects in Italian can have different antecedent preferences in the same context.

[3] There are some additional possible anaphoric subjects that were not taken into account in this chapter. For example, Italian and Dutch allow for a distinction between strong and weak pronouns and between full pronouns and clitics in some cases. The presented diagram only contains the anaphoric subjects that were used in the experimental studies and is therefore a simplification.

The antecedent in subject position is generally the discourse topic, and as a result null pronouns indicate a topic continuation. In contrast, the use of an overt pronoun for reference to the topical antecedent may be perceived as an inappropriate repetition (Belletti & Guasti, 2015), much like the overuse of full NPs (Gordon et al., 1993). Overt pronouns prefer reference to a non–topical, complementary referent, resulting in a topic shift, and are read faster and understood better in topic shift contexts (Carminati, 2002; replicated by Filiaci, 2010). These antecedent preferences have been confirmed in several experimental studies (Serratrice, 2007; Tsimpli, Sorace, Heycock, & Filiaci, 2004). However, antecedent preferences of pronouns are merely biases that can vary on the basis of, for example, implicit causality of the verb (Garvey & Caramazza, 1974), interference of competitor antecedents (Badecker & Straub, 2002; Clackson et al., 2011), and rhetorical relations (Kehler, 2002; Kehler, Kertz, Rohde, & Elman, 2008).

As discussed above, Italian has several types of anaphoric subjects, which, at least some of the time, take different antecedents. However, it is unknown how exactly these anaphoric subjects are processed, and how their meaning is determined. We hypothesize that Italian null subjects, which often refer to the discourse topic, are unmarked. Overt subject pronouns in Italian would then be the marked form, signaling a topic shift.

Because we expect overt pronouns in Italian to be marked, their interpretation is dependent on the meaning of the unmarked null pronoun. As discussed in the introduction, when encountering a marked form such as an overt pronoun, listeners have to reason about why an overt pronoun was used by the speaker, rather than an unmarked null pronoun (Hendriks, 2014). This additional reasoning step of considering the perspective of the speaker when interpreting pronouns is argued to be effortful and to give rise to errors in child language (e.g., De Hoop & Kramer, 2006; Hendriks & Spenader, 2006).

In practice, performing this additional step would mean that listeners first determine the interpretation of the unmarked null pronoun in the presented discourse. This unmarked interpretation is then blocked for an overt pronoun, because if the speaker had wanted to express the unmarked interpretation, a null pronoun would have been used. The result of this additional step of reasoning about an alternative form and blocking its interpretation, is the final interpretation of an overt pronoun. In the case of Italian overt pronouns, a topic continuation interpretation is preferably expressed by a null pronoun. Hence, this interpretation is blocked for an overt pronoun and the topic shift interpretation is the interpretation that remains.

We predict that, if processing a marked overt pronoun in Italian requires reasoning about the alternative unmarked null pronoun, overt pronouns are more effortful to process than null pronouns.

Additionally, if the processing of anaphoric subjects in Italian is indeed characterized by reasoning about alternative forms, then it is expected that in languages in which there are fewer alternative forms, anaphoric subjects are processed differently. This would be the case for non–null subject languages, which have overt pronouns, but no null pronouns. We predict that in non–null subject languages the processing and

interpretation of overt subject pronouns does not depend on reasoning about alternative forms. We will investigate this by comparing the interpretation and processing of anaphoric subjects in Italian to the non–null subject language Dutch. The data on Dutch are described in detail in Vogelzang et al. (2016; see also Chapter 3 of this dissertation), and will briefly be discussed below.

4.1.2 Anaphoric subject processing in Dutch

In their study on anaphoric processing in Dutch, Vogelzang et al. (2016; see also Chapter 3 of this dissertation) found that overt pronouns are interpreted as referring to the discourse topic (in line with Vonk, Hustinx, & Simons, 1992). The interpretation and processing of overt pronouns was compared to the interpretation and processing of full NPs, that unambiguously referred to the discourse topic. Interestingly, overt pronouns were less frequently interpreted as referring to the topic than unambiguous full NPs, indicating that although listeners have a very strong tendency to interpret an overt pronoun as referring to the discourse topic and the most prominent referent (93%), there is still some ambiguity and non–topical interpretations are sometimes allowed. Moreover, interpretation questions about Dutch overt pronouns were answered slower than questions about full NPs.

In addition to the offline measures of final interpretation and response time, Vogelzang and colleagues used the online measure of pupil dilation during the auditory presentation of the stimuli. Pupil dilation can serve as a measure of cognitive effort in language processing (e.g., Beatty & Lucero–Wagoner, 2000; Engelhardt, Ferreira, & Patsenko, 2010; Hyönä, Tommola, & Alaja, 1995; Just & Carpenter, 1993; Scheepers & Crocker, 2004; Schmidtke, 2014; Zellin, Pannekamp, Toepel, & Van der Meer, 2011) and can provide information about the processing of sentences, even without an explicit response from the participant, or provide signatures of cognitive processes that cannot be derived from overt responses (e.g., Van Rijn, Dalenberg, Borst, & Sprenger, 2012; Wierda, Van Rijn, Taatgen, & Martens, 2012). Based on pupil dilation measures, Vogelzang and colleagues found that overt pronouns require more cognitive effort to be processed than full NPs.

Given these results in Dutch, a language that only has overt pronouns, we will investigate how the availability of null pronouns in Italian affects the processing of anaphoric subjects.

4.1.3 Comparison between Italian and Dutch

For the hypothesized correspondence between the referential preferences of anaphoric subjects in Italian and Dutch there are four logical possibilities, which are illustrated in Figure 4.2. These possibilities are presented as four individual hypotheses below.

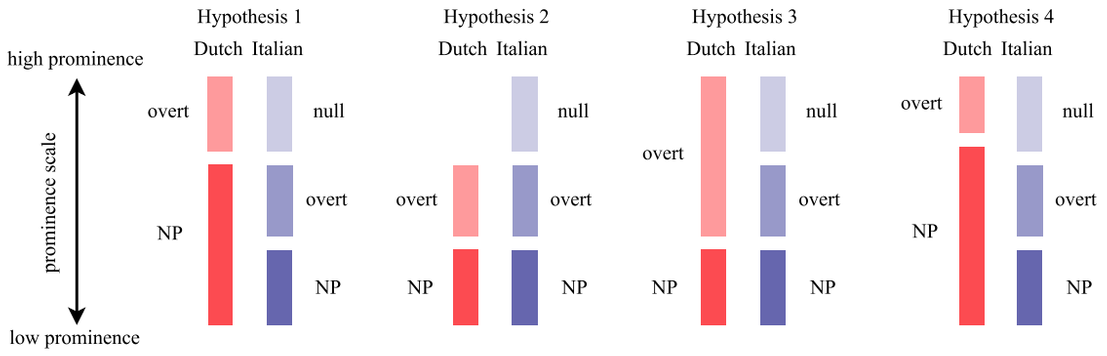


Figure 4.2. The four logical possibilities for the correspondence between the forms of anaphoric subjects in Dutch and Italian (NP: full NP, overt: overt pronoun, null: null pronoun) along a prominence scale ranging from high prominence referents (up) to low prominence referents (down).

The first two hypotheses assume that the meaning of Dutch overt pronouns corresponds directly to the meaning of Italian null or overt pronouns. According to the first hypothesis, Italian *null* pronouns and their meanings correspond to Dutch *overt* pronouns and their meanings. According to the second hypothesis, Italian *overt* pronouns and their meanings correspond to Dutch *overt* pronouns and their meanings, although a further category is present in Italian. Considering this, it is also conceivable that Italian and Dutch pronominal forms do not express exactly the same meanings. This possibility is worked out in the two last hypotheses. If there is no direct correspondence between the meanings of Italian and Dutch pronominal forms, this would mean that one form is more specific in its meaning than another. According to the third hypothesis, Italian overt and null pronouns have a more specific meaning than Dutch overt pronouns. According to the fourth hypothesis, Italian overt and null pronouns have a less specific meaning than Dutch overt pronouns.

4.1.3.1 Hypothesis 1: Italian null pronouns are similar to Dutch overt pronouns

One possibility is that Dutch unstressed overt pronouns and Italian null pronouns are processed and interpreted in a similar way, and differently from Italian overt pronouns. As mentioned, both Dutch overt pronouns and Italian null pronouns generally refer to the discourse topic of the previous sentence, which is the most prominent referent. This suggests that the two forms express the same meaning, and would thus be processed in a similar way. There is little prior research on the comparison of online processing between Italian and Dutch pronouns, but this hypothesis is in line with the claims of Carminati (2002) and Luján (1985) that Italian null pronouns correspond to English unstressed overt pronouns.

This hypothesis predicts that overt pronouns in Italian would be more effortful to process than both null pronouns in Italian and overt pronouns in Dutch because they indicate a topic shift, and topic shifts are harder to process (Van Rij et al., 2013).

4.1.3.2 Hypothesis 2: Italian overt pronouns are similar to Dutch overt pronouns

The second possibility is that Italian overt pronouns correspond to Dutch overt pronouns. These referring expressions are similar in their surface form, as they are both expressed overtly, and so they may be processed in a similar way. According to this second hypothesis, reference to a highly prominent entity cannot be expressed in Dutch through the choice of a particular anaphoric subject, because Dutch only allows for expletive null subjects in very specific situations (Huang, 2000). Consequently, Dutch would not have a subject form that marks high prominence (that is, high prominence should be expressed in some other way in the language). This phenomenon of meanings that cannot be expressed by a particular surface form is known as ineffability (cf. Pesetsky, 1997).

This hypothesis predicts that both Italian and Dutch overt pronouns are marked forms (which are effortful to process), and they indicate a topic shift equally often. In the study of Vogelzang et al. (2016; see also Chapter 3 of this dissertation) it was found that Dutch overt pronouns are interpreted as referring to the discourse topic 93% of the time, so it seems that they are not often interpreted as a topic shift.

4.1.3.3 Hypothesis 3: Italian pronouns are more specific than Dutch overt pronouns

The third possibility reflects the assumption that Dutch anaphoric subjects have a less specific meaning than Italian anaphoric subjects, because Italian has an additional anaphoric subject available and each of these subjects thus refers to a more specific meaning. These more specific anaphoric subjects would be more effortful to process, because listeners would need to reason about alternative forms and their meanings.

This third hypothesis predicts that overt pronouns and null pronouns in Italian refer to those referents that are referred to with an overt pronoun in Dutch. This means that Dutch overt pronouns could express both a topic continuation and a topic shift, and thus that Dutch pronouns have a less specific meaning than Italian pronouns. However, in the study of Vogelzang et al. (2016; see also Chapter 3 of this dissertation) it was found that Dutch overt pronouns are not often interpreted as a topic shift.

4.1.3.4 Hypothesis 4: Italian pronouns are less specific than Dutch overt pronouns

The fourth possibility also assumes that listeners must reason about alternative forms when interpreting anaphoric subjects for which obvious alternatives are available. As a fourth possibility, the meaning of Dutch overt pronouns could be different from the meanings of both Italian overt pronouns and Italian null pronouns.

Because a listener would have to reason about alternative forms, this hypothesis predicts that Italian pronouns will be more effortful to process than Dutch pronouns. As a consequence of Dutch overt pronouns being more specific than Italian pronouns, some meanings that can be expressed by a pronoun in Italian may be expressed by a full NP in Dutch. If this hypothesis holds, the distribution of both Italian anaphoric subjects on the prominence scale will be different from the distribution of the Dutch anaphoric subjects.

In sum, Hypothesis 1 predicts that Italian null and Dutch overt pronouns indicate a reference to the topic equally often, whereas Italian overt pronouns indicate a topic–

shift. Hypothesis 2 predicts that Italian overt and Dutch overt pronouns indicate a reference to the non–topic equally often, whereas Italian null pronouns refer to the topic. Hypothesis 3 predicts that Italian null pronouns refer to the topic, Italian overt pronouns refer to a non–topical referent, and Dutch overt pronouns refer to the topic or the non–topical referent. Hypothesis 4 predicts that Italian null and Dutch overt pronouns both indicate a reference to the topic, but Dutch overt pronouns do so more often than Italian null pronouns. Italian overt pronouns indicate a topic–shift.

4.2 Current study

In this study, we will examine the processing of anaphoric subjects in Italian. By means of pupil dilation measures, we will examine the cost of online processing of different kinds of anaphoric subjects: non–pronominal full NPs such as *the hedgehog*, the overt pronouns *lui* 'he' and *lei* 'she', and null pronouns. We will compare the results of our Italian study to the results from the Dutch study by Vogelzang et al. (2016; see also Chapter 3 of this dissertation). Both experiments recorded responses, response times, and pupil dilation during a referent selection task. Two sets of hypotheses will be tested in this study.

First, Italian null pronouns are the briefest anaphoric subject forms and generally indicate a topic continuation (Carminati, 2002). Therefore, we expect null pronouns to be the unmarked form. Subsequently, overt pronouns would be the marked form, provided that they involve a shift in reference and refer to something other than the discourse topic. If a null pronoun is indeed unmarked and indicates a continuation of the topic, we expect less cognitive effort to be needed to resolve unmarked null pronouns compared to marked overt pronouns, which involve a reference–shift (Carminati, 2002, 2005). Increased cognitive effort would be reflected by inconsistent answers on the referent selection task, longer response times, and increased pupillary responses. Finally, unambiguous full NPs are expected to be easier to process than ambiguous pronouns, because, contrary to ambiguous pronouns, no resolution is needed for unambiguous NPs (cf. Vogelzang et al., 2016; see also Chapter 3 of this dissertation).

Second, a comparison will be made between our findings on the interpretation of anaphoric subjects in Italian and earlier findings on Dutch. Our experiment on Italian uses translated versions of the stimuli of the Dutch study of Vogelzang et al. (2016; see also Chapter 3 of this dissertation), to allow for an optimal comparison between anaphoric interpretations in the two languages. Across the two experiments, similar participant populations and methods were used. Any differences between the Italian experiment and the experiment of Vogelzang and colleagues will be explicitly mentioned in the following sections. We will compare the processing of anaphoric subjects in Italian with the Dutch results on the basis of the four previously discussed hypotheses.

4.3 Methods

4.3.1 Participants

Forty students from the University of Milano–Bicocca, Italy (12 men; mean age 23.5; SD = 5.8) participated in the experiment for course credits. All participants were monolingual native speakers of Italian. All participants had normal or corrected–to–normal vision and hearing.

4.3.2 Materials

The experiment used 96 Italian versions of the same stories as used in Vogelzang et al. (2016; see also Chapter 3 of this dissertation) in Dutch. Whereas Vogelzang and colleagues used 120 stories, time constraints required us to remove 24 stories, which we randomly selected. In addition to anaphoric subjects, the Dutch experiment also tested the interpretation of various anaphoric objects. In order to make parallel versions of the Dutch stimuli, these object manipulations were also included in the Italian stimuli. Only the subset of the data pertaining to anaphoric subjects is of interest for this chapter, but all analyses will be checked for the potential influence of the object manipulations.

A similar story structure was created for all experimental items, in order to provide a coherent discourse in which the anaphoric subject had to be resolved. An example story is presented in Table 4.1. Each story featured two referents, which were displayed as pictures on the screen. The first referent was introduced in the first clause as the grammatical subject. This referent was mentioned as the subject again in the second clause, making it the unambiguous discourse topic. In the second clause a second referent was introduced as the direct or indirect object. Thus, one prominent referent and one less prominent referent were present at this point in the story. The final clause of each story started with the temporal conjunction *mentre* 'while' (cf. e.g., Serratrice, 2007; Sorace & Filiaci, 2006), which introduces a subordinate clause without creating a causal bias. The subject of this third clause was one of three anaphoric subjects: a full NP (e.g., *il riccio* 'the hedgehog'), an overt pronoun (either *lui* ('he') or *lei* ('she')), or a null pronoun (Ø). Two object manipulations were used in the stories: the object was either the reflexive *si* ('himself/herself') or one of the clitic pronouns *lo* ('him') or *la* ('her'). Because the final clause was a subordinate clause, the sentence structure was similar in Italian and Dutch, with the subject and the object preceding the verb. The gender of the pronouns always matched both referents in the story, so gender was not a discriminating cue.

It is important to note that, throughout the entire story exemplified in Table 4.1, *the hedgehog* is the topic on the basis of prominence, frequency, and subject bias, whereas *the mouse* constitutes a non–topical referent. After each story, a referent selection question was asked to determine the interpretation of the anaphoric subject in the third clause. Questions corresponding to the story in Table 4.1 are shown in Table 4.2. If the selected antecedent of a subject question was the same as the subject of the first two

clauses of the story, which is the discourse topic, this was coded as indicating a topic continuation. If, on the other hand, the anaphoric subject was interpreted as referring to the referent that was not the subject of the first two clauses, this was coded as a topic shift (see Figure 4.3 for details on the presentation of the stimuli).

Subject form		**Clause**				
		Il riccio	va con l'aereo in Inghilterra.			
		De egel	gaat met het vliegtuig naar Engeland.			
		The hedgehog	*goes with the airplane to England.*			
		In precedenza	il riccio ha chiesto		al topo	che ora fosse in Inghilterra,
		Eerder	vroeg de egel		aan de muis	hoe laat het was in Engeland,
		Earlier	*the hedgehog asked*		*to the mouse*	*what time it was in England,*
NP		mentre	**il riccio**	**si**	affrettava	verso l'aereo.
		terwijl	**de egel**	**zich**	haastte	naar het vliegtuig.
		while	***the hedgehog***	***himself***	*hurried*	*to the airplane.*
NP		mentre	**il riccio**	**lo**	seguiva	verso l'aereo.
		terwijl	**de egel**	**hem**	volgde	naar het vliegtuig.
		while	***the hedgehog***	***him***	*followed*	*to the airplane.*
overt		mentre	**lui**	**si**	affrettava	verso l'aereo.
		terwijl	**hij**	**zich**	haastte	naar het vliegtuig.
		while	***he***	***himself***	*hurried*	*to the airplane.*
overt		mentre	**lui**	**lo**	seguiva	verso l'aereo.
		terwijl	**hij**	**hem**	volgde	naar het vliegtuig.
		while	***he***	***him***	*followed*	*to the airplane.*
null		mentre	**Ø**	**si**	affrettava	verso l'aereo.
		while	**Ø**	***himself***	*hurried*	*to the airplane.*
null		mentre	**Ø**	**lo**	seguiva	verso l'aereo.
		while	**Ø**	***him***	*followed*	*to the airplane.*

Table 4.1. Example of an experimental story in Italian and Dutch (with English glosses in italics)[4]. All six possible final clauses of the story are listed, sorted by anaphoric subject. The anaphoric subjects investigated in Italian are NP (full NP), overt (overt pronoun), and null (null pronoun). In this story *the hedgehog* is the topic, and *the mouse* is a non–topical referent.

[4] The complete list of stimuli used in the experiment can be found on http://let.webhosting.rug.nl/~vogelzang/experiments.html or by contacting one of the authors.

Object form	Object	Question	Topic continuation answer	Topic shift answer
reflexive	si	Chi si affrettava?	Il riccio	Il topo
	zich	Wie haastte zich?	De egel	De muis
	himself	*Who himself hurried?*	*The hedgehog*	*The mouse*
pronoun	lo	Chi seguiva qualcuno?	Il riccio	Il topo
	hem	Wie volgde iemand?	De egel	De muis
	him	*Who followed someone?*	*The hedgehog*	*The mouse*

Table 4.2. Questions following the story in Table 4.1 in Italian, together with Dutch versions and English glosses. For the different anaphoric subjects, the same questions could be used. For the different objects, different questions were used. Following the story in Table 4.1, the answer *the hedgehog* is coded as a topic continuation, and the answer *the mouse* is coded as a topic shift.

4.3.3 Design

The experiment used a within–subjects design with anaphoric subject as the condition (full NP, overt pronoun, and null pronoun). The experiment consisted of 96 items in total, distributed over four blocks of 24 items with breaks of approximately three minutes in between blocks. The subset of the data with questions about the interpretation of the anaphoric subjects, which will be used for the current analysis, consisted of 48 items (16 critical items per condition). The additional 48 items in the experiment used questions about the interpretation of the object, and may therefore be considered as filler items. Before the main experiment, participants received a practice block of six stories. The complete experiment took approximately 60 minutes. Like in Vogelzang et al. (2016; see also Chapter 3 of this dissertation), anaphoric subject and the location of the target referent on the screen (left or right) were counterbalanced across trials. The order of presentation of the stories was randomized per participant.

4.3.4 Procedure

Following the same procedure as Vogelzang et al. (2016; see also Chapter 3 of this dissertation), participants were tested individually using an EyeLink 1000 (SR Research) eye tracker to measure pupil dilation and a Microsoft Sidewinder gamepad to record responses. The gamepad was used for accurate measures of response times and so that participants could keep looking at the screen while giving their responses. Participants heard the stories and questions through headphones while looking at the computer screen. The participants were instructed to listen carefully and to blink as little as possible during trials. To reduce the number of blinks during trials, participants were explicitly asked to blink after every trial. Between blocks, the eye tracker was re–

calibrated.

Throughout the duration of each trial, the two referents in each story were presented as pictures on the screen; one on the left side and the other on the right side. To prevent the pictures from influencing the pupillary measures during the test sentence, they appeared before the beginning of the story and remained visible until after the question had been answered. Participants could answer the referent selection questions by selecting one of the referents on the screen through the left or right trigger button on the back of the gamepad. After the question was answered, a black frame was displayed around the selected referent. No other feedback was provided. Figure 4.3 shows the structure of presentation of the stimuli during a trial.

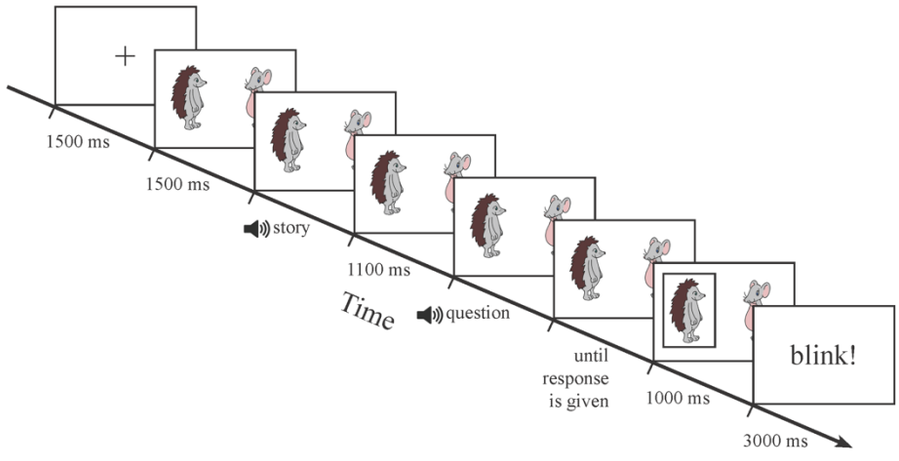


Figure 4.3. The structure of presentation of the stimuli during one trial. The story started 1500 ms after presentation of the referents. 1100 ms after the story, the question was presented. The participant could respond by selecting one of two referents on the screen. Figure taken from Vogelzang et al. (2016; see also Chapter 3 of this dissertation).

4.4 Results

In the time window ranging from 1000 ms before the onset of the word following the critical anaphoric subject until the onset of the question, samples recorded during blinks were encoded as missing data. Trials that contained more than 25% missing data in this time window were excluded. The missing data of the remaining trials were interpolated. In total, 198 trials (10.3%) were excluded and 1722 trials remained for analysis. For all analyses, we collapsed over object manipulations.

For all comparisons with Italian we used both the full Dutch dataset with 120 stories and the subset of the Dutch data with the same 96 stories that were used in the Italian study. No qualitative differences were found between comparisons of the Italian data to the full dataset and the subset of the data for Dutch. Therefore, the Dutch data with 120 stories that was presented in Vogelzang et al. (2016; see also Chapter 3 of this dissertation) will be reported on in this chapter.

The results of the offline responses will be presented first, followed by the response times (RTs) and the pupil dilation data.

4.4.1 Responses

4.4.1.1 Italian

The responses to the questions in the three anaphoric subject conditions for Italian are shown in Figure 4.4a. The responses are expressed as the percentage of responses that indicate a continuation of the topic.

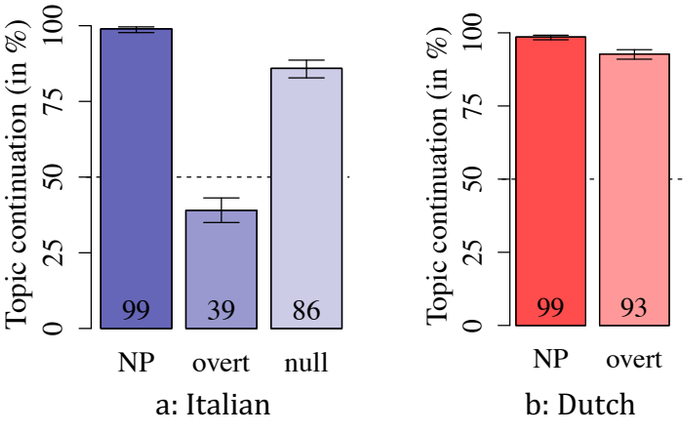


Figure 4.4. The percentage of responses indicating a topic continuation for the three anaphoric subjects in Italian (NP: full NP, overt: overt pronoun, null: null pronoun) are presented in Figure 4.4a. The percentage of responses indicating a topic continuation for the two anaphoric subjects in Dutch (NP: full NP, overt: overt pronoun) are presented in Figure 4.4b. Error bars are derived from logistic analysis.

The responses to the interpretation questions indicate that in Italian, unambiguous full NPs were consistently interpreted correctly as referring to the discourse topic (99%). As expected, null pronouns were generally interpreted as referring to the discourse topic as well (86%). For overt pronouns, there were both topic continuation interpretations (39%) and topic shift interpretations (61%).

These responses were analyzed using a binomial generalized linear mixed effect–based model. Based on the theoretical predictions anaphoric subject was included as a fixed effect. Following Baayen, Davidson, and Bates (2008), we entered both participants and items as random intercepts and random slopes and assessed, by means of model comparisons, whether the inclusion of these random factors was warranted. The best model included both random intercepts for participants and for items. In line with recommendations of Barr, Levy, Scheepers, and Tily (2013) we also ran a model with the maximal random effects structure. The results from this model showed the same effects as the model with reduced random effects structure. The results from the latter model are reported below.

A comparison of the full model including an effect of anaphoric subject against a model excluding the effect of anaphoric subject revealed an effect of anaphoric subject ($\chi^2(2) = 681.28$; $p < 0.001$). The full model including anaphoric subject indicated that participants gave more topic continuation responses with full NPs than with overt pronouns ($\beta = 5.39$; $z = 12.43$; $p < 0.001$) or null pronouns ($\beta = 2.84$; $z = 6.57$; $p < 0.001$). Also, participants selected more topic continuation responses with null pronouns than with overt pronouns ($\beta = 2.55$; $z = 15.02$; $p < 0.001$).

A Welch two–sample t–test showed that the antecedent preference for overt pronouns (61%) significantly deviated from 50%, and thus that overt pronouns show a preference towards a non–topic antecedent ($t(39) = 3.43$; $p < 0.01$). A planned comparison showed that the topic preference for null pronouns was stronger than the non–topic preference for overt pronouns ($\beta = 1.40$; $z = 9.32$; $p < 0.001$). Importantly, all the reported effects were present with both object manipulations.

4.4.1.2 Comparison between Italian and Dutch

The results from the Dutch study are shown in Figure 4.4b. Planned comparisons with generalized linear mixed–effects models showed no differences in the interpretation of full NPs between Italian and Dutch (99% vs. 99%, $\beta = 0.27$; $z = 0.47$; $p = 0.64$). A comparison of Dutch overt pronouns (93% topic continuation) and Italian null pronouns (86% topic continuation) revealed that Dutch overt pronouns have a stronger preference for the topical antecedent than Italian null pronouns ($\beta = 0.84$; $z = 3.54$; $p < 0.001$).

4.4.2 Response times

4.4.2.1 Italian

The response times (RTs) to the questions were measured from the end of the question until the first button press. The mean RT per condition for Italian is shown in Figure 4.5a.

The RTs were analyzed using a linear mixed effect–based model of the log–transformed RTs. Based on the theoretical predictions, anaphoric subject was included as a fixed effect. Based on model comparisons (cf. Baayen et al., 2008), random intercepts for participants and random slopes for anaphoric subjects within participants were included. In line with recommendations of Barr, Levy, Scheepers, and Tily (2013) we also ran a model with the maximal random effects structure. The results from this model showed the same effects as the model with reduced random effects structure. The results from the latter model are reported below.

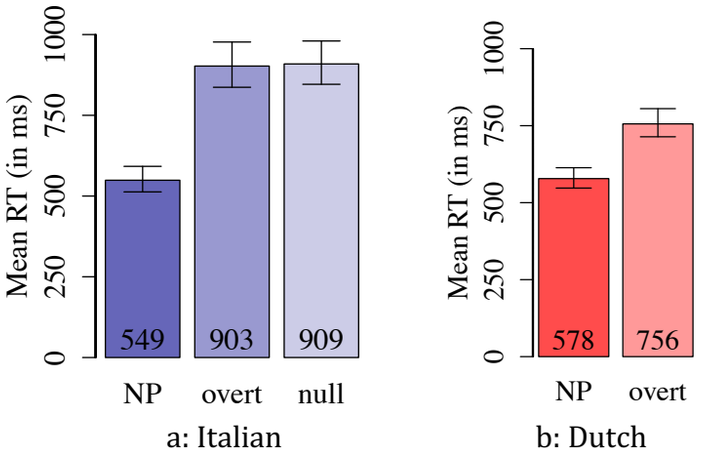


Figure 4.5. The mean response times for the interpretation questions regarding the three anaphoric subjects in Italian (NP: full NP, overt: overt pronoun, null: null pronoun) are presented in Figure 4.5a. The mean response times for the interpretation questions regarding the two anaphoric subjects in Dutch (NP: full NP, overt: overt pronoun) are presented in Figure 4.5b. Error bars are based on a 95% two–sided nonparametric confidence interval estimated through bootstrapping.

A comparison of the full model including an effect of anaphoric subject against a model excluding the effect of anaphoric subject revealed an effect of anaphoric subject ($\chi^2(2) = 28.39$; $p < 0.001$). The full model including anaphoric subject indicated that full NPs are responded to faster than overt pronouns ($\beta = 0.41$; $t(35) = 5.75$; $p < 0.001$) and null pronouns ($\beta = 0.45$; $t(35) = 5.83$; $p < 0.001$). No difference in RT was found between overt and null pronouns ($\beta = 0.04$; $t(37) = 0.59$; $p = 0.57$). Importantly, all the reported effects were present with both object manipulations.

4.4.2.2 Comparison between Italian and Dutch

The results from the Dutch study are shown in Figure 4.5b. When comparing the RTs of the questions in Italian and Dutch, planned comparisons with generalized linear mixed–effects models showed no difference between the response times to Italian and Dutch full NPs (resp. 549 ms and 578 ms, $\beta = 0.03$; $t(79) = 0.37$; $p = 0.712$). Additional comparisons showed that neither the difference between the response times to Italian overt pronouns and Dutch overt pronouns (resp. 903 ms vs. 756 ms, $\beta = 0.17$; $t(86) = 1.59$; $p = 0.117$) nor the difference between the response times to Italian null pronouns and Dutch overt pronouns was significant (909 ms vs. 756 ms, $\beta = 0.21$; $t(86) = 1.93$; $p = 0.057$).

4.4.3 Pupil dilation data

Before further analysis, the pupil dilation data were down–sampled to a sampling rate of 100 Hz. For the analysis of the pupil dilation data, only the time window from 1000 ms before the onset of the word following the critical anaphoric subject in the third clause

until the onset of the question was taken into account. Because the time window did not include the referent selection questions, the full data set was used for the pupil dilation analysis. After excluding trials that contained more than 25% missing pupil dilation data, 3490 trials remained for analysis.

4.4.3.1 Italian

The pupil dilation data for each anaphoric subject over the course of a trial, averaged over all trials, are plotted in Figure 4.6. The dilation is expressed as proportional change relative to a baseline window. This baseline was determined by first calculating the mean pupil dilation for each trial separately, from 100 ms before the onset of the subject until the onset of the subject. The samples from each trial were subsequently divided by their corresponding baseline in order to get the proportional pupil dilation. A positive proportional pupil dilation (above zero) indicates that the pupil size, interpreted as reflecting processing effort, increased during the test sentence.

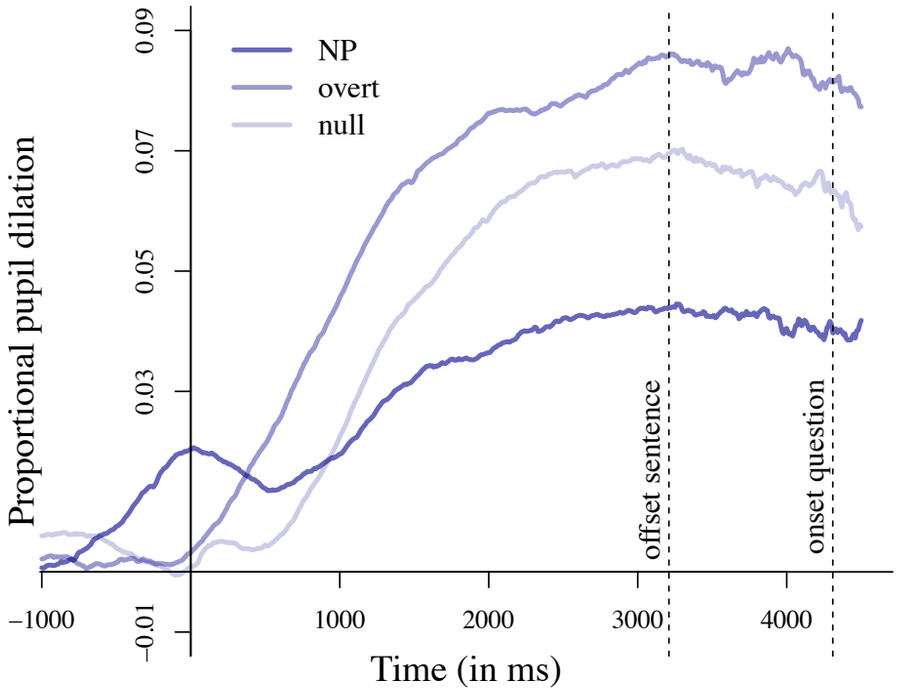


Figure 4.6. The proportional pupil dilation over time for the three anaphoric subjects in Italian (NP: full NP, overt: overt pronoun, null: null pronoun). The time is aligned to the onset of the word following the subject. Dotted lines indicate the means of the offset of the test sentence and the onset of the question. A baseline window was calculated for each trial. Note that the baseline window varies per trial and cannot be easily identified in the graph. See text for detailed information on the alignment and baselining of the data.

The dilation data were aligned to the onset of the object that followed the anaphoric subject. This was done because the anaphoric subjects in the experiment varied in length (e.g., *the hedgehog* vs. ∅). The mean onset of the anaphoric subjects in the plot

therefore varies from –1000 ms for full NPs to –500 ms for overt pronouns, and to 0 ms for null pronouns (for which the onset of the subject and the object was taken to be at the same time point). Thus, full NPs already start around 1000 ms before t=0 in Figure 4.6, and therefore the NP condition shows a rise in pupil dilation before t=0.

We used Generalized Additive Models (GAMs; Wood, 2006) to analyze the pupil dilation data. This allows for an analysis of the data over time for all conditions using one model. Based on the theoretical predictions, the GAMs model used anaphoric subjects as a fixed effect. Based on model comparisons, random slopes and intercepts for participants, random slopes and intercepts for items, an effect of trial order, and an effect of time over centered trials were included.

The results of the model indicate a difference in pupil dilation between sentences with overt and null pronouns compared to full NPs (resp. $F(1,17.90) = 63.43$, $p < 0.001$ and $F(1,17.12) = 22.00$, $p < 0.001$). Recall that the onsets of the different anaphoric subjects vary; this explains why the dilation in sentences with overt pronouns starts rising earlier than with null pronouns in Figure 4.6. From around 1000 ms after the onset of the word following the subject, both the conditions with overt pronouns and with null pronouns show more dilation than full NPs. From 2000 ms on, the pupil dilation stabilizes, with overt pronouns eliciting the most pupil dilation, before null pronouns and full NPs, which elicit the least pupil dilation. An additional analysis indeed shows that the pupil dilation curve for sentences with overt pronouns is different from the curve with null pronouns ($F(1,16.92) = 25.88$, $p < 0.001$).

4.4.3.2 Comparison between Italian and Dutch

When comparing the Italian results to the Dutch results, it is not possible to compare the absolute pupil dilation of the two experiments. Pupil dilation is a measure that is very sensitive to lighting conditions in the lab, but also to other influences such as attention and possibly language–specific influences, and so it is not possible to compare the proportional dilation of the two experiments directly. It is possible to test participants under seemingly more comparable conditions in the same lab, but nevertheless factors such as testing in a first or foreign language context could influence this result[5].

Figure 4.7 presents the pupil dilation data from the Dutch study of Vogelzang et al. (2016; see also Chapter 3 of this dissertation). The difference in amplitude between the Italian pupil dilation data (max. amplitude around 0.09) and Dutch (max. amplitude around 0.03) is clear, and is most likely caused by differences in lighting at the different test locations. However, a comparison between the *relative* effects found in the two experiments is possible.

[5] Pupil dilation is a measure that is very sensitive to internal as well as external influences. To check for influence of the specific location in which the participants were tested, we conducted a pilot experiment with the same setup as the described experiments, testing Italian participants (n=10) and Dutch participants (n=11) in the same lab in the Netherlands. The results showed pupil dilation effects similar to the ones reported in this chapter, but the amplitudes of the relative pupil dilation were much more similar between Italian and Dutch in this pilot study that in the main experiment. However, we chose to test Italian participants in their home country for the main experiment in order to limit foreign language exposure contamination.

In Dutch, it was found that sentences with overt pronouns elicit more pupil dilation than sentences with full NPs. This is in line with the Italian results, where both sentences with overt pronouns and sentences with null pronouns elicit more pupil dilation than sentences with full NPs.

However, there are also differences between the Italian and the Dutch results. In Italian, the effects of processing the clause are already visible before the onset of the word following the anaphoric subject (in the case of an NP) or right after the onset of the word following the anaphoric subject (in the case of an overt pronoun). In Dutch, the rise in pupil dilation in sentences with full NPs and in sentences with overt pronouns starts much later, around 1000 ms.

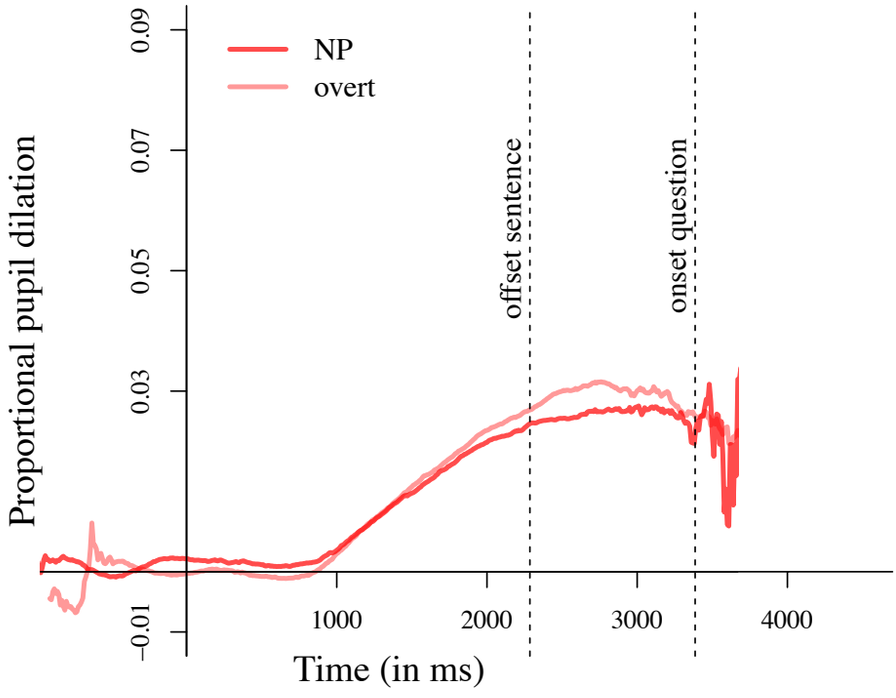


Figure 4.7. The proportional pupil dilation over time for the two anaphoric subjects in Dutch (NP: full NP, overt: overt pronoun). The time is aligned to the onset of the word following the subject. Dotted lines indicate the means of the offset of the test sentence and the onset of the question.

We assessed whether the increase in pupil dilation after the onset of the word following the subject differed as a function of language. It should be noted that this analysis is post–hoc, as we did not expect to see any effects in slope based on theoretical considerations. To eliminate the difference in maximum amplitude between the Italian and Dutch datasets, we normalized all pupil dilation data to the mean of the amplitude between the offset of the sentence and the onset of the question as assessed in the NP condition. We calculated regression slopes of the pupillary response for the first 1000 ms after the onset of the word following the subject for each participant and for each condition separately for both Italian and Dutch. A post–hoc linear mixed effect–based

model showed that the average slope of the Italian data is steeper than the slope of the Dutch data ($\beta = 0.39$; $t = 3.76$; $p < 0.001$).

4.5 Discussion: A cross–linguistic comparison

In this chapter, we described a referent selection study that investigated the interpretation and processing of anaphoric subjects in Italian, a null subject language. We further compared the results of this study to results from an experiment in Dutch, a non–null subject language. Our referent selection task measured interpretation, response time, and pupil dilation with sentences containing anaphoric subjects. The results of the Italian study will be discussed first, and will then be compared to the results of the Dutch study of Vogelzang et al. (2016; see also Chapter 3 of this dissertation).

4.5.1 Italian

We hypothesized that overt pronouns would be the marked form in Italian, and would require the listener to take an additional step of reasoning about an alternative form. This additional reasoning step would be reflected by less consistent answers on the referent selection task, longer response times, and increased pupillary responses. Full NPs were expected to be easier to process than pronouns.

As predicted, the results show effects of the form of anaphoric subject in Italian, with unambiguous full NPs being interpreted more often as referring to the discourse topic than potentially ambiguous pronouns.

In the introduction, we suggested that null pronouns are the unmarked form, and overt pronouns are the marked form (cf. Horn's division of pragmatic labor, 1984). Our results indeed show a division of pragmatic labor, because the different anaphoric subjects have different meanings. In line with Carminati (2002), null pronouns were generally interpreted as signaling a topic continuation, and overt pronouns were most often interpreted as signaling a topic shift in our experiment. In other words, null pronouns generally refer to referents with a high prominence, and overt pronouns generally refer to referents with a lower prominence.

However, the results also show that the two Italian subject pronouns do not always have clearly distinct meanings; null pronouns refer to the discourse topic 86% of the time and overt pronouns refer to the same topic 39% of the time. Thus, at least part of the time these subject forms are interpreted as referring to the same entity. This indicates that Horn's division of pragmatic labor with respect to anaphoric subjects in Italian is a gradient rather than absolute distinction.

When looking at the response times and the pupil dilation measures, full NPs are responded to faster, and elicit less pupil dilation than potentially ambiguous pronouns. These results are in line with our predictions about the processing of full NPs compared to pronouns, and indicate that pupil dilation is a viable measure of processing effort.

The difference in processing between overt pronouns and null pronouns was examined through both response times and pupil dilation. Although we do not find any differences between these two pronominal forms in the response times, we do in the pupil dilation results. Overt pronouns elicit more pupil dilation than null pronouns. This indicates that the differences in processing effort that are reflected in pupil size during the sentence were no longer present or measurable after the referent selection question. Thus, these results show that the measure of pupil dilation is more sensitive than the measure of response times as a reflection of online processing difficulty.

Overall, the Italian results show differences in interpretation as well as online processing between the different anaphoric subjects, with more specific anaphoric subjects being easier to process than less specific anaphoric subjects. We argue that the difficulties in the interpretation and the processing of the marked overt pronoun stem from the execution of a costly additional reasoning step, which is necessary to reason about alternative forms (cf. Hendriks, 2014).

4.5.2 Comparison between Italian and Dutch

The results of Vogelzang et al. (2016; see also Chapter 3 of this dissertation) showed that Dutch overt pronouns preferably refer to a topical antecedent. This was also found for Italian null pronouns in the present study. However, the antecedent preference of Dutch overt pronouns is stronger than that of Italian null pronouns. There are no significant differences in response times between Italian and Dutch pronouns. Unambiguous full NPs are interpreted quickly and correctly in both languages.

In the introduction, we presented four hypotheses regarding the relation between Italian and Dutch anaphoric subjects. The weaker antecedent preference found for Italian null pronouns suggests that they are not processed in the same way as Dutch overt pronouns, in spite of the fact that they both have a preference for topic continuation (contra Hypothesis 1). Moreover, it shows that Dutch overt pronouns have a more specific meaning than both Italian overt pronouns and Italian null pronouns, so there are meanings that can be expressed by a pronoun in Italian that cannot be expressed by a pronoun in Dutch (contra Hypotheses 2 and 3). The finding that Dutch pronouns indicate a topic continuation more often than Italian null pronouns provides evidence for Hypothesis 4 (see Figure 4.2).

As discussed in the introduction, the difference in meaning between anaphoric subjects in Italian and Dutch may be related to the availability of different anaphoric subjects in these languages[6]. When a language offers a choice between a marked and an unmarked form (cf. Horn, 1984), as in Italian, as opposed to the availability of a single

[6] When comparing anaphoric subjects in different languages, it is unavoidable to test these subjects in combination with other aspects of the languages. Thus, there are other factors that could influence the observed processing differences, such as pragmatic, syntactic, morphological, and historical factors. The experiments were kept as similar as possible in design and materials in order to allow for a direct comparison between Italian and Dutch, but care must still be taken when reasoning about the cause of the observed differences.

form, as in Dutch, the choice of a specific form provides additional information about the intended referent. This may restrict the possible meanings for the form (cf. Kiparsky, 1982). This additional information has to be taken into account when resolving the anaphoric subject and leads to a specific meaning for the pronominal form.

In the pupil dilation data analyses, it was found that Italian sentences with overt pronouns and null pronouns show more pupil dilation than sentences with full NPs. Dutch sentences with overt pronouns also showed more pupil dilation than those with full NPs. Although the pupil dilation effects cannot be directly compared across the Italian and Dutch studies, the effect of pronouns being costlier to process than full NPs is apparent in both languages. This is in line with the offline analyses of the interpretations and RTs, showing that unambiguous full NPs are easier to process and interpret than pronouns across languages.

Notably, two seemingly conflicting conclusions were drawn in this chapter about the processing difficulty of more informative anaphoric subjects. On the one hand, full NPs that provide unambiguous information about the intended referent are easier to process in both Italian and Dutch than pronouns, that need to be resolved. On the other hand, overt and null pronouns in Italian, for which the specific form used provides additional information about the intended meaning, are argued to be more difficult to process than Dutch pronouns. The crucial difference between these two cases is the availability of alternative forms. When encountering an unambiguous full NP, a listener does not need to reason about alternative forms. The processing and interpretation will be easy in this case. On the other hand, when encountering a potentially ambiguous form like an overt pronoun in Italian, the listener will need to reason about why that form was used instead of the alternative form of a null pronoun. Thus, whereas in general more informative or less ambiguous forms are easier to process (full NPs easier than pronouns, Italian null pronouns easier than Italian overt pronouns), on the basis of our cross–linguistic comparison we argue that processing difficulty is also influenced by the alternative forms available in a language to express similar meanings.

An unexpected finding was the difference in onset of the pupillary effects over time. The pupillary responses in Dutch were delayed compared to Italian. Although such an effect was not included in our initial hypotheses, the earlier effects for Italian could indicate that Italian anaphoric subjects are more informative than Dutch anaphoric subjects, and therefore elicit an earlier pupillary increase. Further research on the timing of pupillary responses (such as the work of Wierda, Van Rijn, Taatgen, & Martens (2012) using a temporal attention task) is necessary to be able to fully interpret delayed effects.

4.5.3 General discussion

We examined the processing of different anaphoric subjects, and compared this between two languages. However, the results can be generalized beyond the specific scope of this study. Three general points will be discussed below, being 1) pupil dilation being a sensitive measure of processing effort, 2) how the results can be extended to the

processing of expressions other than anaphoric subjects, and 3) how words that have similar but slightly different meanings can be categorized.

Firstly, the study showed the value of examining processing effort and complexity of conditions during sentence processing by means of pupillometry. Whereas the response times to interpretation questions did not differentiate between Italian overt pronouns and Italian null pronouns, the online measure of pupil dilation showed a clear effect. Additionally, the pupil dilation data revealed differences in the timing of pronoun processing across languages. Thus, pupillary responses are a sensitive measure of online processing effort that can reveal effects that offline measures do not demonstrate.

In this chapter, it was assumed that increased pupillary responses reflect cognitive effort of language processing, as has been suggested in the literature (e.g., Beatty & Lucero–Wagoner, 2000; Engelhardt, Ferreira, & Patsenko, 2010). However, other factors may also have influenced pupil size, such as (implicitly or explicitly) deciding to do something (cf. Katidioti, Borst, & Taatgen, 2014), like moving the eyes towards a referent, or interpreting a referring expression in a certain way. In our study, all these potential influences are expected to cause a larger pupil size when a referent is more difficult to process. So, whether the measured pupil size reflects referent processing directly or reflects processes related to referent processing, the differences in pupil size found in the experiment can still be interpreted as reflecting cognitive effort. However, more research on pupillary responses is surely needed to better understand the many factors influencing pupil dilation.

Secondly, the finding that overt Italian pronouns are more effortful to process than null pronouns can be extended to the processing of other expressions that have seemingly similar meanings, such as anaphoric objects. Italian distinguishes between clitics (e.g., *lo* 'him') and strong or stressed forms (e.g., *lui* 'him'). On the basis of the results presented in this chapter, we would predict that anaphoric objects also show a division of labor between an unmarked and a marked form, resulting in differences in processing effort.

This prediction of processing differences for expressions that have similar meanings also has possible implications for the acquisition and use of the mental lexicon. When multiple expressions are available that seemingly refer to the same entity, like 'coach' and 'bus', the processing of the expression could be influenced by how specific the expression is. In other words, expressions that are assumed to be synonyms in certain contexts, may not be equally easy to process. To investigate this issue, many studies that have examined synonymous expressions could be repeated, now measuring the processing difficulty by means of pupil dilation.

Finally, the study shows that different, seemingly similar forms in different languages can have different meanings. Words with a similar phonology or orthography in different languages, but with different meanings, have previously been named *false friends* (Aronoff & Rees–Miller, 2002). Examples of a false friend are the Italian 'attuale', which in English means 'present', but not 'actual' (Ferguson, 1994), and the Italian 'caldo', which means 'hot', not the phonologically and orthographically more similar 'cold'. The anaphoric subjects used in our cross–linguistic comparison suggest another

category, namely that of syntactically similar forms with different meanings, which we will call *fake friends.* Fake friends, unlike false friends, do not necessarily have a similar phonology or orthography. Nonetheless, they might lead to some confusion because they have the same syntactic category, and this, in turn, might lead to the wrong suggestion that the two forms have the same meaning. In fact, their meanings differ in a subtle way. In particular, the two forms interact differently with their linguistic context. Thus, whereas false friends are words that have no semantic overlap, fake friends have a large, but not complete, semantic overlap. We argue that Dutch overt pronouns and Italian overt and null pronouns do not adhere to the classical definition of false friends, but can be seen as fake friends.

To conclude, we investigated how the availability of alternative referential forms in a language affects the meaning and processing of a form. The results show that different anaphoric subjects are processed and interpreted differently in Italian. Moreover, different anaphoric subjects are processed and interpreted differently in the null subject language Italian compared to the non–null subject language Dutch. We argue that the availability of alternative forms affects the processing of a form (see also Sprenger & Van Rijn, 2013). This is based on the assumption that an additional reasoning step is needed for the processing of forms for which competing alternatives are available.